(12) UK Patent Application (19) GB (11) 2 352 915 (13) A

(43) Date of A Publication 07.02.2001

(21) Application No 9918646.2

(22) Date of Filing 06.08.1999

(71) Applicant(s)
Television Monitoring Services Limited
(Incorporated in the United Kingdom)
9 Gloucester Place, WITNEY, Oxon, OX8 5LA,
United Kingdom

(72) Inventor(s)
Cameron Hay

(74) Agent and/or Address for Service
Mathys & Squire
100 Grays Inn Road, LONDON, WC1X 8AL,
United Kingdom

(51) INT CL^7
G06K 9/20

(52) UK CL (Edition S)
H4F FGXX F3P F3T
G1A AA3 ABGX AG6 AG9 AR7 AT26 AT27 AT3 AT4

(56) Documents Cited
EP 0893779 A1 EP 0720114 A2 EP 0569311 A1
US 5867593 A

(58) Field of Search
UK CL (Edition R) G1A AAJL AAJX ABAX ABGX
AMSX , H4F FGXD FGXX
INT CL^7 G06K 9/20
Online: WPI, EPODOC, JAPIO

(54) Abstract Title

A method of retrieving text data from a broadcast image

(57) Text data is retrieved from a broadcast image by capturing a sequence of frames at predetermined intervals. Each captured frame image is processed to emphasise text content and de-emphasise changing backgrounds materials. The presences or absence of text in the frame is detected by measuring a luminance ratio of the brightest part of the image (the potential text) and the background part. Selected frames are then passed to OCR software for text retrieval.

Figure 2E

GB 2 352 915 A

Figure 1

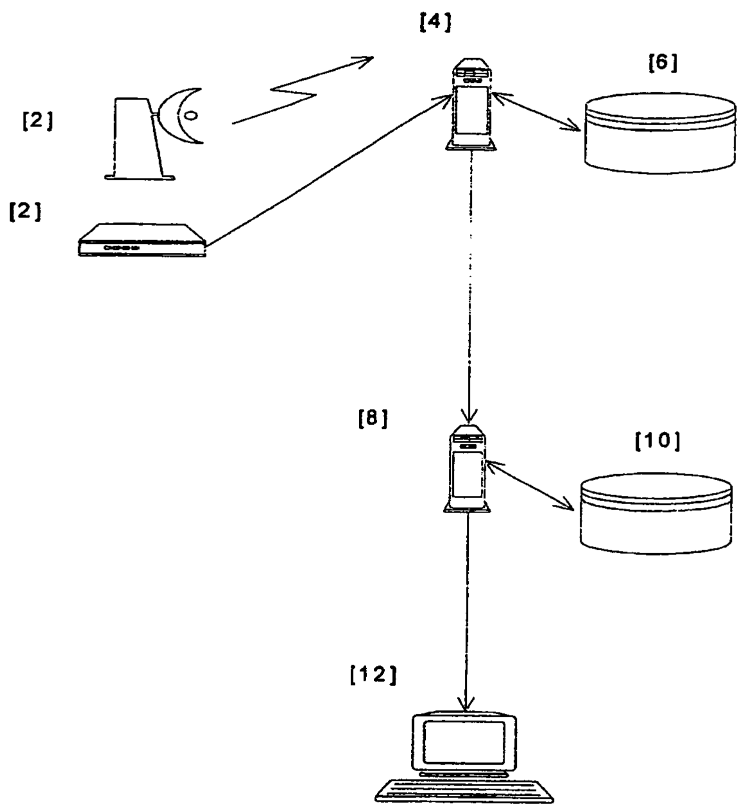

Figure 2A

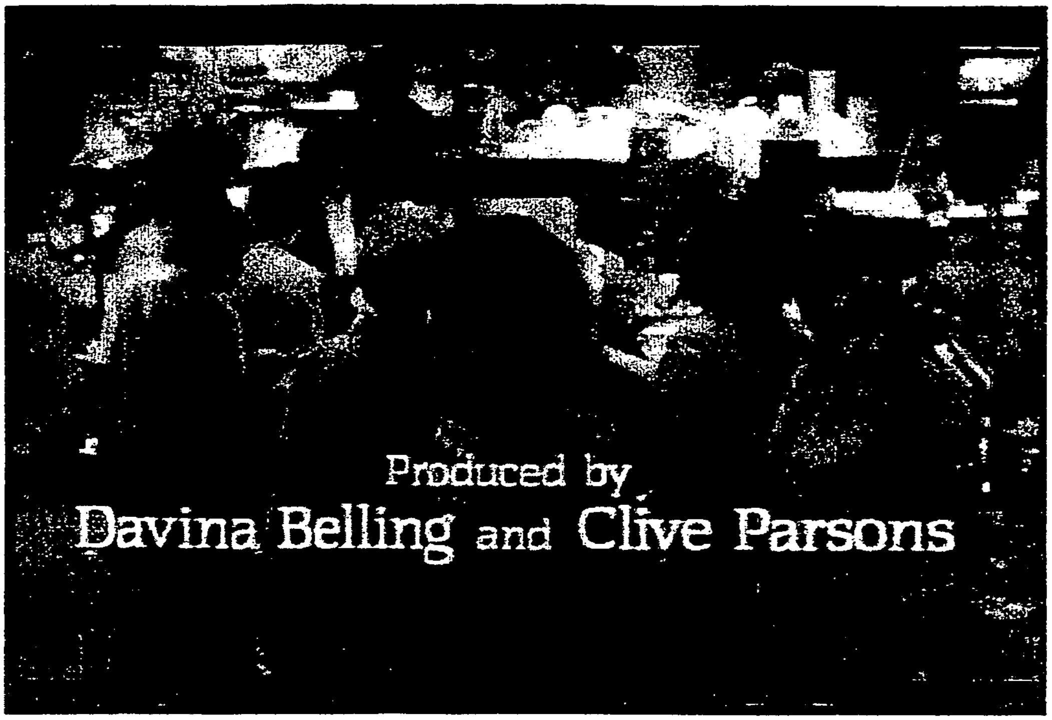

Figure 2B

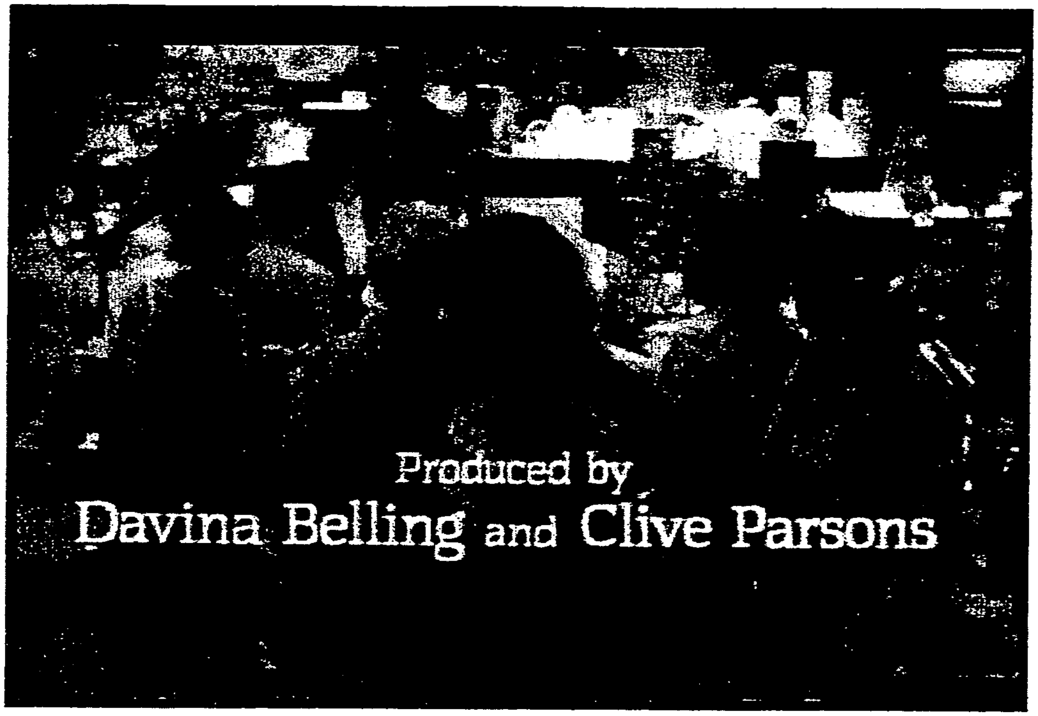

Figure 2C

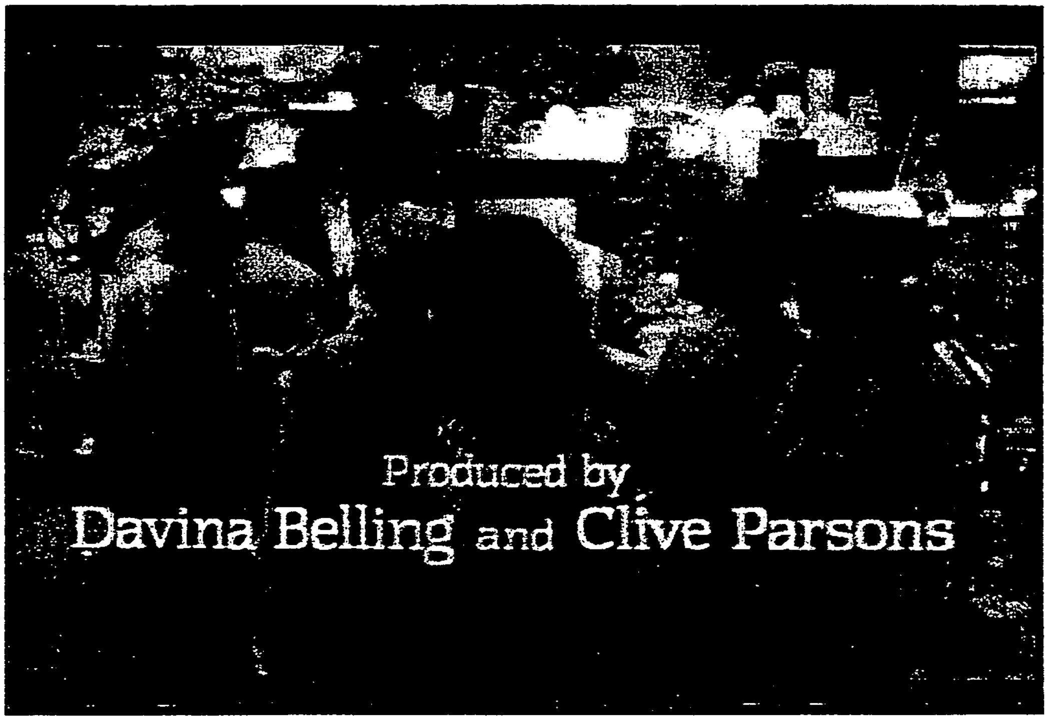

Figure 2D

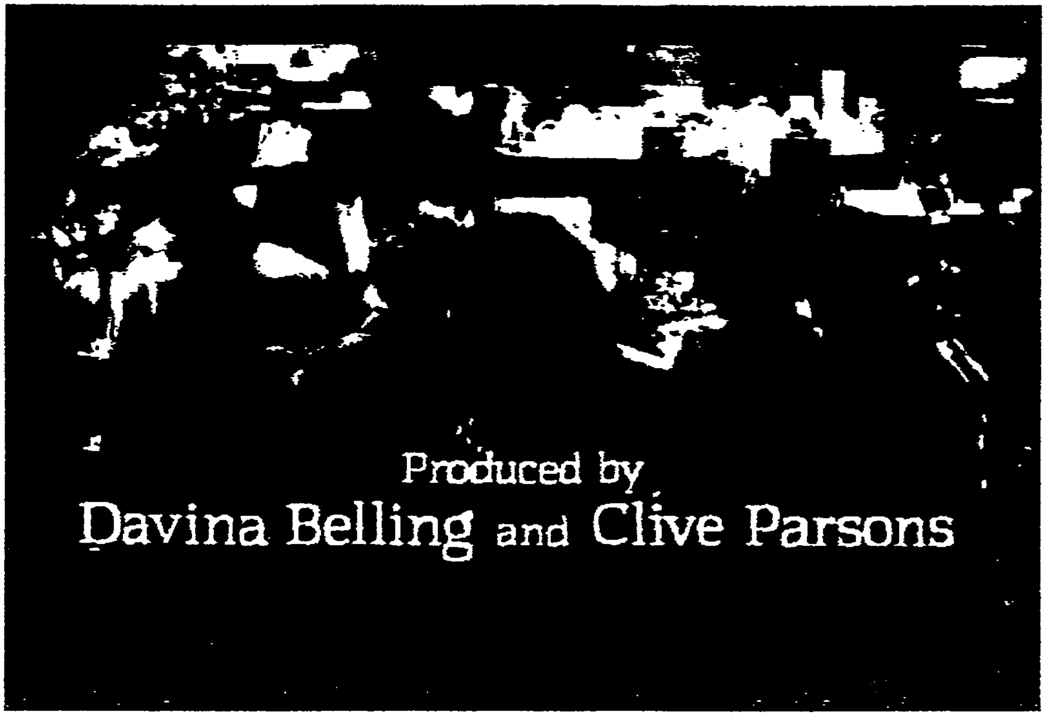

Figure 2E

2352915

A method of retrieving text data from a broadcast image

Background of the Invention

The present invention relates to a method of retrieving text data from a broadcast image, and, more particularly, to the reading of information conveyed in text within the credits of a broadcast program.

Technical Problem

Currently a large number of different institutions around the world monitor television broadcasts for various purposes. In many cases there are several institutions in each country performing the same task. Generally this is done manually. With the increasing numbers of channels broadcast the task is expensive and unreliable.

The reasons for monitoring broadcasts are numerous. For example programme producers may wish to verify where and when their programmes were broadcast in order to check copyright returns from broadcasters. The information is also required for audience research.

Prior art solutions to improve on simply watching and noting down details of programs have been limited to scrolling through time lapse video recordings of the channels in order to limit the volume of images recorded and to reduce the time needed to monitor a channel.

Solution of the Invention

The present invention provides a method of retrieving text data from a broadcast image; comprising the steps of capturing a sequence of frame images from a source at predetermined intervals, processing the frame images,

measuring a ratio of the luminescence of a brightest part of the processed frame image relative to a background part of the processed frame image,
selecting the processed frame images where the ratio exceeds a predetermined threshold,
processing the selected images with OCR software to read text in the selected images.

This solution takes advantage of the fact that in order for text to be legible to the viewer it is generally enhanced relative to the background by giving it much greater luminescence than the rest of the image. A simple luminescence ratio test on the processed frame image can therefore be used to identify frames containing text for storage and further OCR processing. As the number of frames containing text is far less than the remainder of the non-text images, the process of the present invention can retrieve text data from a broadcast image in "effective" real time.

The retrieved text is preferably stored in a database. The integrity of the data can be enhanced by spell checking it against a database of known titles and personal names of those involved in the production business. Since such a system will read with a high accuracy all the text data that is broadcast for a programme, a database of all the information contained in the rolling credits can be created. This data with the data relating to the channel monitored and time information is invaluable in copyright and contract enforcement work as well as in many other applications.

Brief Description of the Drawings

In order that the invention may be well understood an embodiment thereof will now be described, by way of example only, with reference to the accompanying diagrammatic drawings, in which:

Figure 1 shows a block diagram of a processing system for retrieving data from a broadcast image; and

Figure 2A-2E show a series of sample frames created using the process of the present invention.

A source 2 provides an output, which is the form of a video signal. The figure shows two alternative sources 2 to illustrate that the source may be derived from a variety of video stream types such as directly off air (terrestrial analogue and digital signals, satellite and cable transmissions) or may be taken from videotape onto which the broadcast has been recorded for subsequent processing, or it may be the output of a decoder in the case of digital or pay channels.

The present invention is also suitable for use with a DVD (digital video disc) source as long as the output play back is through a standard SCART, RF or S-Video connector in order to provide a video signal. The stored signal on a DVD is already compressed as an MPEG2 – a compression technique which stores the changing pixels from one frame to the next. Therefore it is not possible to store a single frame image to process if the content of the DVD are not played back as mentioned.

A frame grabber 4 is connected to the output of the source 2. The frame grabber is set to sample the source at a rate selected so that any text data will show in at least one frame. A rate of one frame per second would normally ensure that several frames of any relevant text will be captured. Preferably the system allows a predetermined interval between frame captures to be set by the user depending on the circumstances of the application. For example if the system is to be used to retrieve subliminal text broadcast for such a short period that it is below the threshold of viewer perception, the predetermined interval must be sufficiently short to ensure that one of the few frames containing the text to be retrieved is captured.

Suitable frame grabbers include a SNAPPER (Trade Mark) board which contains circuitry to convert the captured frame image into a digital image compressed to the

JPEG standard. A BRONCO (Trade Mark) frame grabber that outputs a bitmap image could also be used.

The captured frame images are passed to a memory 6 for buffer storage. At this stage the typical size of a single frame image will be about 1.2MB.

A processor 8 then processes the captured frame images. The processor can be the CPU of a PC.

One processing step is to convert the captured frame image to a black and white image.

A technical problem that this processing is designed to overcome is the disturbance to text characters caused by the interlaced raster scans of a television picture. Since a captured frame image will include parts of two separate interlaces the processor separates the interlaces by taking an average of the adjacent lines, which come from the separate interlaces. The further processing is carried out on the averaged value only.

Additional processing steps may be carried out to identify and remove background information from the image. The term background in this context means image data that appears behind text data usually at a lower luminosity. This background material is liable to change more rapidly from frame to frame than the text elements of the captured frame image. This aids identification of background material for elimination.

The objective is to prepare an image that can be inverted to produce a black on white image. The image can also be compressed with any suitable compression algorithm to reduce the size of the stored frame. Typically after removing background material the size of a processed frame image will be of the order of 20k making the storage of large quantities of text containing frames feasible.

The luminosity of the brightest part of the processed frame image is then compared to

the average luminosity of the processed frame image to derive a luminosity ratio for the processed frame image. If the measured ratio is greater than a preset threshold the frame is passed to a permanent store 10. If the ratio is below the threshold the frame is discarded. A monitor 12 may be provided to display the processed frame image and this can be used to set the threshold adaptively by user intervention if needed. Certain automated dynamic threshold changes may also be made. After the described processing images which contain text data will show text clearly whereas the picture frames will have been reduced to a uniform texture.

The next step is to process the selected images with OCR software to derive the text therein. The same processor 8 can be used for this purpose. Various OCR products are available such as TEXTBRIDGE ® from XEROX ® that will take as input a bitmap image and create an output text file. The text file can then be imported into a database and subjected to various data clean up techniques. Since the data is rolling credits from broadcast programmes, spell checking against a database of known titles and personal names will produce significant improvement in data quality.

Figure 2 shows sample frame images taken at various stages in the process. Figure 2A shows the original colour frame as captured by the frame grabber 4. Figure 2B shows the image after conversion to 256 shades of grey. Figure 2C shows the same frame after conversion to 16 shades of grey. Figure 2D shows the image in a "dropout" form with only two tones. The luminescence threshold for converting to one tone or the other is set to exclude as much of the non text data as possible. Here there were several bright points of the background image. This image is then inverted to give the black on white image which is Figure 2E. This image can be fed to an OCR program which will readily recognise the words shown and also deliver some random characters, which are easily recognised as such. Spell checking and the training facilities provided with most OCR software correct the errors of recognition to leave just the text data required as the output of the process.

It has been found that the accuracy of the process when applied to television production images is close to 100% with slightly reduced accuracy when identifying data from the credits of feature films since these are normally produced for cinema and are consequently broadcast in a smaller size. Titles of programs can normally be read with good accuracy. Poor performance will be encountered in the few situations where the text data of the credits has been made legible by means other than heightened luminosity. This is unusual. In some cases however hard edge techniques may be used to enable text to be picked out by the viewer from a background of the same luminosity.

Since only a proportion of the captured frame images needs to be selected, the process can produce text output virtually instantaneously. Therefore the process can be used to switch the video stream from the source 2. For example, if a broadcaster wishes to divert the broadcasts of news or weather programmes to an Internet site the detection of the text titles of these programmes by the processor 8 can be used to control a switch (not shown) to divert the video stream. Similarly the detection of certain text data may be used to trigger selective recording of a broadcast. For example characteristic text data at breaks in programmes can be used to record the advertisements or stop a recording process so that they are eliminated.

It will be appreciated that the real time retrieval of text data from a broadcast image has many useful applications that will be apparent to the man skilled in the art.

Claims

1. A method of retrieving text data from a broadcast image; comprising the steps of capturing a sequence of frame images from a source at predetermined intervals, processing the frame images,
measuring a ratio of the luminescence of a brightest part of the processed frame image relative to a background part of the processed frame image,
selecting processed frame images where the ratio exceeds a predetermined threshold,
processing the selected images with OCR software to read text in the selected images.

2. A method as claimed in claim 1, wherein the text is stored.

3. A method as claimed in claim 1, wherein the broadcast image is recorded by a video recorder prior to the capturing step.

4. A method as claimed in claim 1, wherein the processing step comprises averaging the pixel values in adjacent lines of a raster scan within the image to eliminate the effect of interlacing.

5. A method as claimed in claim 1, wherein the selected images are stored and the remaining images are discarded.

6. A method as claimed in claim 1, wherein the frames are captured at the rate of one per second.

7. A method as claimed in claim 1, wherein the predetermined threshold is adaptively set.

8. A method of retrieving text data from a broadcast image substantially as herein described with reference to the accompanying drawing.

The Patent Office

INVESTOR IN PEOPLE

9.

Application No: GB 9918646.2
Claims searched: All

Examiner: Joe McCann
Date of search: 28 February 2000

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.R): G1A(AMSX, AAJL, AAJX, ABAX, ABGX); H4F(FGXD, FGXX)

Int Cl (Ed.7): G06K(9/20)

Other: Online: WPI, EPODOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
Y	EP 0893779A1	(SONY) - see abstract	1-3,6
Y	EP 0720114A2	(SIEMENS) - see abstract	1-3,6
Y	EP 0569311A1	(IBM) - see abstract	1-3,6
Y	US 5867593A	(TATSUTA ET AL) - see abstract	1-3,6

X Document indicating lack of novelty or inventive step
Y Document indicating lack of inventive step if combined with one or more other documents of same category.
& Member of the same patent family

A Document indicating technological background and/or state of the art.
P Document published on or after the declared priority date but before the filing date of this invention.
E Patent document published on or after, but with priority date earlier than, the filing date of this application.